Procedia Computer Science
Volume 51, 2015, Pages 1791–1800
ICCS 2015 International Conference On Computational Science

Total Least Squares and Chebyshev Norm

Milan Hladík[1] and Michal Černý[2]

[1] Charles University, Department of Applied Mathematics, Prague, Czech Republic
milan.hladik@matfyz.cz
[2] University of Economics, Department of Econometrics, Prague, Czech Republic
cernym@vse.cz

Abstract
We investigate the total least square problem (TLS) with Chebyshev norm instead of the traditionally used Frobenius norm. The use of Chebyshev norm is motivated by the need for robust solutions. In order to solve the problem, we introduce interval computation and use many of the results obtained there. We show that the problem we are tackling is NP-hard in general, but it becomes polynomial in the case of a fixed number of regressors. This is the most important practical result since usually we work with regression models with a low number of regression parameters (compared to the number of observations). We present not only a precise algorithm for the problem, but also a computationally efficient heuristic. We illustrate the behavior of our method in a particular probabilistic setup by a simulation study.

Keywords: Total least squares, Chebyshev norm, interval computation, computational complexity

1 Introduction and motivation

Notation. For a vector $v \in \mathbb{R}^n$ and a matrix $M \in \mathbb{R}^{m\times n}$, we use vector L_p norms $\|v\|_1 = \sum_i |v_i|$, $\|v\|_2 = \sqrt{\sum_i v_i^2}$, $\|v\|_\infty = \max_i |v_i|$, and matrix norms: Frobenius norm $\|M\|_F = \sqrt{\sum_{i,j} M_{ij}^2}$ and Chebyshev (max) norm $\|M\|_{\max} = \max_{i,j} |M_{ij}|$. Next, $|M|$ denotes the entrywise absolute value and M_{i*} the i-th row of a M. Eventually, I_n stands for the identity matrix of size n, and E and e for the matrix and the vector of ones (with suitable dimensions), respectively.

Motivation. We often need to work with data suffering from imprecision, instability, uncertainty or various kinds of errors. For example, in econometrics we often work with estimated future values suffering from prediction errors; in numerical analysis, data are affected by rounding errors; in operations research, data often represent planned values (such as flight times or processing times) which differ from the true ones. Many further examples of erroneous data can be found in other areas as well.

This paper is a contribution to a particular problem in regression analysis where erroneous data can appear both in the output variable and the input variables.

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2015

doi:10.1016/j.procs.2015.05.393

OLS, TLS and problem formulation. In linear regression and data fitting, we try to find x that "suits best" a given overdetermined system of equations $Ax = b$. In case of ordinary least squares (OLS) it is assumed that errors are in b and we are to find the closest b' such that $Ax = b'$ is consistent, where the distance between b and b' is measured by the L_2-norm. It is well-known that under some circumstances, other norms are also useful; e.g. in robust statistics one often minimizes $\|b - b'\|_1$ or $\|b - b'\|_\infty$. Another example of usage of a norm different from L_2 is the case when errors follow the Laplace distribution; then the minimization of $\|b - b'\|_1$ leads to maximum likelihood estimation.

Total least-squares (TLS) is a natural generalization of OLS admitting errors in both A and b; see [5, 12, 22]. The usual problem formulation is: find $(A' \mid b')$ such that $A'x = b'$ is consistent and $\|(A \mid b) - (A' \mid b')\|_F$ is minimal. The above mentioned ideas from statistics, leading to the replacement of L_2-norm by another norm, can be reformulated to the TLS case in a straightforward way. In general, we are interested in solving the TLS problem when the Frobenius norm is replaced by another norm. (This problem is much less studied than TLS, see [10, 13, 17, 22, 23, 24].) Said otherwise, the general question is:

$$\begin{array}{l}\text{given an overdetermined system } Ax = b \text{ and a matrix norm } \|\cdot\|, \text{ find } (A' \mid b') \\ \text{such that } A'x = b' \text{ is solvable and } \|(A \mid b) - (A' \mid b')\| \text{ is minimal.}\end{array} \tag{1}$$

Motivation also arises from the fact that the classical TLS solutions are sometimes ill-conditioned, even more than OLS solutions [4, 12]. Thus, searching for an appropriate norm in (1) yielding more robust solutions is of high importance.

Our contribution. This paper is a contribution to the general question (1) for the case of Chebyshev norm.

Remark 1. Observe that the problem has also applications in numerical analysis. Say that we know that a system $A'x = b'$ has a solution, but the exact data A', b' of the system are not available; what is available is an imprecisely computed matrix A and an imprecisely computed vector b. Let each element of $(A \mid b)$ differ from the corresponding element in $(A' \mid b')$ by at most $\delta \geq 0$, where δ is assumed to be as small as possible. Then, solving (1) with the Chebyshev norm can be seen as a method of reconstruction of $(A' \mid b')$ from $(A \mid b)$, overcoming the problem of the imprecise computation.

Interval computation. The main tool of our analysis is theory of computation with interval vectors and matrices. By an interval matrix we mean a family of matrices

$$\boldsymbol{A} := [\underline{A}, \overline{A}] = \{A \in \mathbb{R}^{m \times n};\ \underline{A} \leq A \leq \overline{A}\},$$

where $\underline{A} \leq \overline{A}$ are given and inequalities between matrices/vectors are understood entrywise. The set of all $m \times n$ interval matrices is denoted by $\mathbb{IR}^{m \times n}$. Interval vectors are defined and denoted accordingly.

Let $\boldsymbol{A} \in \mathbb{IR}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{IR}^m$. Then the *solution* of the interval system $\boldsymbol{A}x = \boldsymbol{b}$ is any $x \in \mathbb{R}^n$ such that $Ax = b$ for some $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Thus, the *solution set* to $\boldsymbol{A}x = \boldsymbol{b}$ is defined as

$$\{x \in \mathbb{R}^n;\ \exists A \in \boldsymbol{A},\ \exists b \in \boldsymbol{b} : Ax = b\}.$$

Notice that we use this definition not only for square, but also for overdetermined systems.

An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is *regular* if every matrix $A \in \boldsymbol{A}$ is nonsingular; otherwise $\boldsymbol{A}$ is *singular*. Similarly, $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ has *full column rank* if every $A \in \boldsymbol{A}$ has full column rank.

An *enclosure* of a set $\mathcal{S} \subset \mathbb{R}^n$ is any interval vector $\boldsymbol{v} \in \mathbb{IR}^n$ such that $\mathcal{S} \subseteq \boldsymbol{v}$.

One of major problems of interval analysis is to find a tight enclosure to the solution set of a given system of interval-valued linear equations. For some classical methods and recent developments see, e.g., [2, 6, 15, 18]. Overdetermined systems in particular were discussed in [8, 9].

Interval-theoretic formulation of (1). Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. We are to solve the optimization problem

$$\min \|(\Delta A \mid \Delta b)\|_{\max} \ \text{ subject to } \ (A + \Delta A)x = b + \Delta b \text{ is solvable.} \tag{2}$$

We can formulate (2) in terms of interval computation as

$$\min \delta \ \text{ subject to } \ [A - \delta E, A + \delta E]x = [b - \delta e, b + \delta e] \text{ is solvable.} \tag{3}$$

In this paper, we will investigate a more general problem

$$\min \delta \ \text{ subject to } \ [A - \delta A^\Delta, A + \delta A^\Delta]x = [b - \delta b^\Delta, b + \delta b^\Delta] \text{ is solvable,} \tag{4}$$

where $A^\Delta \in \mathbb{R}^{m\times n}$ and $b^\Delta \in \mathbb{R}^m$ are nonnegative.

Remark 2. The fact that we admit a general matrix $A^\Delta \geq 0$ and a general vector $b^\Delta \geq 0$ can give us more flexibility (but the reader could always think of $A^\Delta = E$ and $b^\Delta = e$). We can, for example, set $A^\Delta := |A|$ and $b^\Delta := |b|$, which results in computation of minimal relative perturbations of given data such that the system is solvable.

Remark 3. So far, interval computation was used in linear regression to deal with interval input data [1, 7]. Formulation (4) creates a link between linear regression and interval computation, and opens the door to a possible application of interval techniques in TLS.

Optimal solutions. In the following sections, δ^{opt} denotes the optimal value and x^{opt} denotes an optimal solution of (4).

2 Determining δ^{opt} and x^{opt}

Consider the interval system $\boldsymbol{A}_\delta x = \boldsymbol{b}_\delta$, where

$$\begin{aligned} \boldsymbol{A}_\delta &:= [A - \delta A^\Delta, A + \delta A^\Delta], \\ \boldsymbol{b}_\delta &:= [b - \delta b^\Delta, b + \delta b^\Delta]. \end{aligned}$$

By the Oettli–Prager theorem [2, 16], the solution set $\mathcal{S}$ of this interval system is characterized as

$$\mathcal{S} = \{x \in \mathbb{R}^n : |Ax - b| \leq \delta A^\Delta |x| + \delta b^\Delta\}. \tag{5}$$

This is a nonlinear system, but it can be reformulated by means of 2^n linear systems. Actually, $x \in \mathbb{R}^n$ is a solution iff there is $s \in \{\pm 1\}^n$ such that the linear system

$$(A - \delta A^{\Delta} D_s)x \le b + \delta b^{\Delta}, \tag{6a}$$

$$(-A - \delta A^{\Delta} D_s)x \le -b + \delta b^{\Delta}, \tag{6b}$$

$$D_s x \ge 0 \tag{6c}$$

is solvable, where $D_s = \operatorname{diag}(s)$. Thus, δ^{opt}, x^{opt} can be determined as

$$\delta^{\mathrm{opt}} = \min_{s\in\{\pm 1\}^n} \min\{\delta \ge 0;\ (6) \text{ is solvable}\}. \tag{7}$$

The right-hand side optimization problem of (7) can be written as

$$\min_{s\in\{\pm 1\}^n} \underbrace{\min_{x\in\mathbb{R}^n} \max_{\substack{i\in\{1,\dots,m\}\\ j\in\{0,1\}}} \frac{(-1)^{1-j}A_{i*}x + (-1)^j b_i}{A^{\Delta}_{i*}D_s x + b^{\Delta}_i} \ \text{ subject to }\ A^{\Delta}D_s x + b^{\Delta} \ge 0,\ D_s x \ge 0,.}_{(\star)}$$

The inner optimization problem $(\star)$ is a generalized linear fractional programming problem (GLFP) solvable in polynomial time using e.g. an interior point method [3, 14]. (The constraint $A^{\Delta}D_s x + b^{\Delta} \ge 0$ is obviously redundant, but we have stated it explicitly in order it be clear that $(\star)$ is indeed a GLFP.) We have proved the first important result:

Theorem 1. *Problem (1) can be solved in time $O(2^n \cdot p(\mathrm{bitsize}(A, b)))$, where p is a polynomial and* bitsize *denotes the length of binary representations of rational numbers in A, b.* □

The following corollary captures the case which is important for practice: the point is that we usually have regression models with a low number of regression parameters compared to the number of observations. (Said otherwise, computation time exponential in the number of regression parameters is "good news" — an algorithm exponential in the number of observations would be indeed bad.)

Corollary 1. *When the number of regression parameters is fixed, that is*

$$n = O(1), \tag{8}$$

then problem (1) can be solved in polynomial time. □

Unfortunately, relaxation of the assumption (8) leads to a hardness result, which will be proved in the next section.

2.1 Complexity

We show that computation of δ^{opt} is NP-hard even for the case with $A^{\Delta} = E$, $b^{\Delta} = e$. NP-hardness is proved for a natural decision version of the optimization problem. The hardness result shows that the computation bound of Theorem 1 is in some sense best possible and that we are indeed "lucky" that there exists an algorithm which is exponential in n but not in m.

Theorem 2. *Let $\alpha \ge 0$. Then it is NP-hard to decide whether $\delta^{\mathrm{opt}} \le \alpha$ on a sub-class of problems with $A^{\Delta} = E$, $b^{\Delta} = e$, and $m = n + 1$.*

Proof. First notice that the condition $\delta^{opt} \le \alpha$ is equivalent to solvability of $[A-\alpha E, A+\alpha E]x = [b-\alpha e, b+\alpha e]$.

Now, let $M \in \mathbb{Q}^{n\times n}$ and $\alpha \in \mathbb{Q}$, $\alpha \ge 0$. Checking whether the interval matrix $[M-\alpha E, M+\alpha E]$ is singular is an NP-hard problem [11]. The interval matrix $[M-\alpha E, M+\alpha E]$ is singular iff there is $i \in \{1,\dots,n\}$ and $x^* \in \mathbb{R}^n$, $x^*_i = 1$ such that x^* solves $[M-\alpha E, M+\alpha E]x = 0$. In other words, the interval system

$$[M^{\neg i} - \alpha E, M^{\neg i} + \alpha E]x = [M_{*i} - \alpha e, M_{*i} + \alpha e] \tag{9}$$

has a solution, where M_{*i} denotes the ith column of M and $M^{\neg i}$ denotes the matrix M with the ith column deleted. Thus, if we can decide on solvability of (9) in polynomial time, we could check for singularity of interval matrices in polynomial time as well. Therefore the problem in question is NP-hard. □

2.2 Finding a minimizer

We know how to compute x^{opt} and δ^{opt}. When we are to find a minimizer $(A' \mid b')$ of (1), it suffices to solve the system

$$A'x^{\text{opt}} = b',\ A - \delta^{\text{opt}}A^\Delta \le A' \le A + \delta^{\text{opt}}A^\Delta,\ b - \delta^{\text{opt}}b^\Delta \le b' \le b + \delta^{\text{opt}}b^\Delta \tag{10}$$

with data $\delta^{\text{opt}}, x^{\text{opt}}, A, A^\Delta, b, b^\Delta$ and variables A', b'. This is a linear feasibility problem which can be solved in polynomial time by linear programming techniques. Moreover, (10) describes exactly the set of all minimizers. On the other hand, if we are interested in only one minimizer, we can use the explicit formula from [2]: Define $z := \text{sgn}(x)$ and define the vector $y \in [-1,1]^m$ as

$$y_i = \begin{cases} \dfrac{(Ax^{\text{opt}} - b)_i}{\delta^{\text{opt}}(A^\Delta|x^{\text{opt}}| + b^\Delta)_i} & \text{if } \delta^{\text{opt}}(A^\Delta|x^{\text{opt}}| + b^\Delta)_i > 0, \\ 1 & \text{otherwise.} \end{cases}$$

Now, we can put

$$A' := A - \delta^{\text{opt}} \cdot \text{diag}(y)A^\Delta\,\text{diag}(z), \quad b' := b + \delta^{\text{opt}} \cdot \text{diag}(y)b^\Delta.$$

2.3 Properties

Existence. As in the classical TLS, the optimal solution needn't exist. As an example, consider

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix},\ b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},\ A^\Delta = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix},\ b^\Delta = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Here, the optimal value of (4) is unbounded. Typically, however, the optimal value is bounded. This is the case, for instance, for the most natural choices $A^\Delta = E$, $b^\Delta = e$ or $A^\Delta = |A|$, $b^\Delta = |b|$ since $\delta^{opt} \le \max_{i,j}\{|a_{ij}|, |b_i|\}$ holds in the former and $\delta^{opt} \le 1$ holds in the latter.

Conversely, even when the optimal value is bounded, the optimal value still needn't be attained. For example, when

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},\ b = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ A^\Delta = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ b^\Delta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Here, any $\delta > 0$ is feasible, but $\delta = 0$ is not. Such problems are called *non-generic* in the classical TLS.

Denote

$$\delta^{\text{inf}} := \inf \delta \;\text{ subject to }\; [A - \delta A^{\Delta}, A + \delta A^{\Delta}]x = [b - \delta b^{\Delta}, b + \delta b^{\Delta}] \text{ is solvable.}$$

As long as δ is bounded from above, we have $\delta^{\text{inf}} < \infty$. If $\boldsymbol{A}_{\delta}x = \boldsymbol{b}_{\delta}$ is solvable for $\delta = \delta^{\text{inf}}$, then δ^{opt} is attained and $\delta^{\text{opt}} = \delta^{\text{inf}}$. The following observation gives another condition under which (4) has an optimal solution.

Proposition 1. *Suppose that $\boldsymbol{A}_{\delta}$ has full column rank for $\delta = \delta^{\text{inf}}$. Then $\delta^{\text{opt}} = \delta^{\text{inf}}$.*

Proof. For an interval matrix $\boldsymbol{M} \in \mathbb{IR}^{m\times n}$, denote its smallest singular value as

$$\sigma_{\min}(\boldsymbol{M}) := \min\{\sigma_{\min}(M);\; M \in \boldsymbol{M}\}, \tag{11}$$

where $\sigma_{\min}(M)$ is the smallest singular value of a real-valued matrix M. Due to continuity of $\sigma_{\min}(\cdot)$ for real matrices and compactness of $\boldsymbol{M}$, the minimum in (11) always exists.

Under the assumption of the proposition, $\boldsymbol{A}_{\delta}x = \boldsymbol{b}_{\delta}$ is solvable iff the interval matrix $(\boldsymbol{A}_{\delta} \mid \boldsymbol{b}_{\delta})$ has not full column rank. Equivalently, $\sigma_{\min}(\boldsymbol{A}_{\delta} \mid \boldsymbol{b}_{\delta}) = 0$. Since $\sigma_{\min}(\boldsymbol{A}_{\delta} \mid \boldsymbol{b}_{\delta}) = 0$ for every $\delta > \delta^{\text{inf}}$, it vanishes also for $\delta = \delta^{\text{inf}}$ and therefore δ^{opt} is attained. □

Uniqueness. Similarly as for L_{∞}-regression (or: Chebyshev approximation), we uniqueness of the optimal solution is not guaranteed in general. As an example, consider

$$A = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\; b = \begin{pmatrix} 1 \\ -1 \end{pmatrix},\; A^{\Delta} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},\; b^{\Delta} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

for which $\delta^{\text{opt}} = 1$ and x^{opt} is any real value.

Another formulation. In view of (5), problem (3) can be expressed as

$$\min \delta \;\text{ subject to }\; |Ax - b| \le \delta E|x| + \delta e \text{ is solvable,}$$

which is equivalent to

$$\min_{x} \frac{\|Ax - b\|_{\infty}}{\|x\|_1 + 1}.$$

This gives a nice link to the classical TLS, which can be formulated as optimization problem

$$\min_{x} \frac{\|Ax - b\|_2^2}{\|x\|_2^2 + 1}.$$

3 A heuristic method

The negative result of Theorem 2 shows that we might meet instances for which the computation of δ^{opt} is intractable, even in spite of the good news of Corollary 1. Then we must turn to heuristics. Here we present two useful ideas, which can be used as a basis for more sophisticated methods or combined with general metaheuristic approaches. (We are convinced that design of heuristics for (1) is a tempting topic deserving a separate research.).

3.1 Lower bound on δ^{opt}

The system $Ax = b$ is infeasible if the matrix $(A \mid b)$ has full column rank. Thus, the interval system $\boldsymbol{A}_\delta x = \boldsymbol{b}_\delta$ has no solution if the interval matrix $(\boldsymbol{A}_\delta \mid \boldsymbol{b}_\delta)$ has full column rank, that is, it contains only full column rank matrices.

Full column rank of interval matrices were investigated in several papers; see, e.g., [19, 20, 21]. Among the known methods, the following performs well. An interval matrix $\boldsymbol{M}$ has full column rank if $\rho(|(M^c)^\dagger|M^\Delta) < 1$, where $\rho(\cdot)$ denotes the spectral radius and $(\cdot)^\dagger$ the Moore–Penrose pseudoinverse. Applying this sufficient condition to our case with $\boldsymbol{M} := (\boldsymbol{A}_\delta \mid \boldsymbol{b}_\delta)$, we get the condition

$$\rho\left(|(A \mid b)^\dagger| \cdot \delta \cdot (A^\Delta \mid b^\Delta)\right) = \delta \cdot \rho\left(|(A \mid b)^\dagger| \cdot (A^\Delta \mid b^\Delta)\right) < 1.$$

In other words, if

$$\delta < \rho\left(|(A \mid b)^\dagger| \cdot (A^\Delta \mid b^\Delta)\right)^{-1},$$

then $(\boldsymbol{A}_\delta \mid \boldsymbol{b}_\delta)$ has full column rank and $\boldsymbol{A}_\delta x = \boldsymbol{b}_\delta$ is infeasible. Hence,

$$\delta^{opt} \geq \rho\left(|(A \mid b)^\dagger| \cdot (A^\Delta \mid b^\Delta)\right)^{-1}.$$

3.2 Upper bounds on δ^{opt}

We propose two upper bounds on δ^{opt}.

First, let x^* be the traditional least-squares solution of $Ax = b$, or any other heuristic solution. Denote $s := \mathrm{sgn}(x^*)$. Solve the GLFP $(\star)$ associated with this sign vector s and denote by δ^* its optimal value. Then

$$\delta^{opt} \leq \delta^*.$$

Second, let δ^F be the optimal value of the classical TLS problem

$$\min \|(\Delta A \mid \Delta b)\|_F \ \text{ subject to } \ (A + \Delta A)x = b + \Delta b \text{ is solvable},$$

which can be found easily with SVD decomposition.

Since $\|\cdot\|_{\max} \leq \|\cdot\|_F$, we have

$$\delta^{opt} \leq \delta^F.$$

The drawback of this approach is that it does not give us explicitely the corresponding perturbations ΔA and Δb.

3.3 The resulting heuristic

Now, we present a heuristic for determining δ^{opt}.

Calculate an upper bound $\delta^U \geq \delta^{opt}$ by any method mentioned in Section 3.2. Let $\boldsymbol{x} \in \mathbb{IR}^n$ be an enclosure of the solution set of the overdetermined interval system $\boldsymbol{A}_\delta x = \boldsymbol{b}_\delta$.

If $\boldsymbol{x}$ crosses only a small number of orthants, than we can effectively calculate δ^{opt}, x^{opt} by (7), where s is subject to sign vectors of all crossing orthants instead of $\{\pm 1\}^n$.

In particular, if δ^* is the upper bound from Section 3.2, and if $\boldsymbol{x}$ lies in one orthant only, then $\delta^{opt} = \delta^*$ and x^{opt} is the corresponding optimal solution of the GLFP $(\star)$.

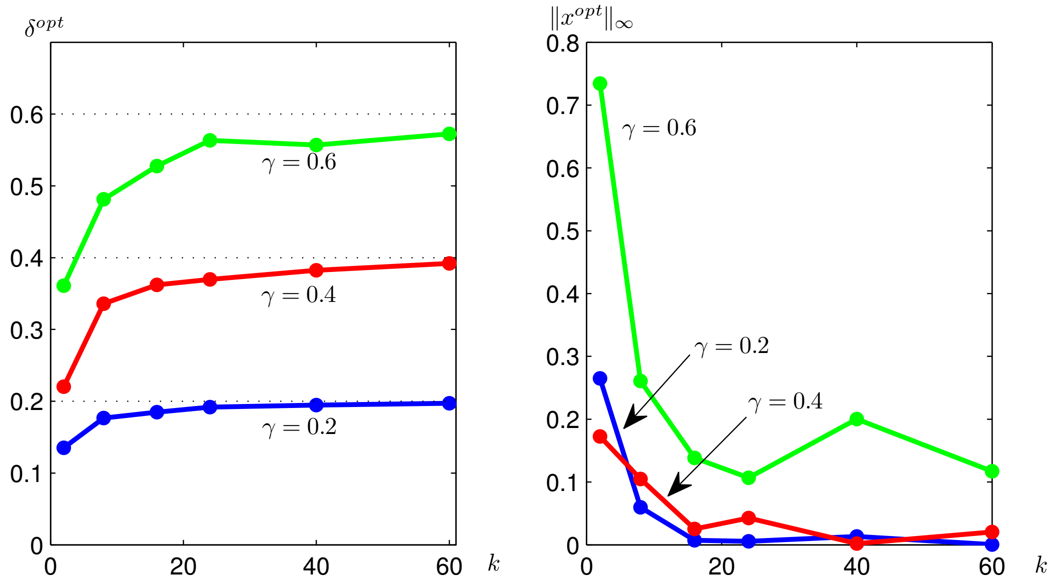

Figure 1: Average simulated values of δ^{opt} (left plot) and $\|x^{\mathrm{opt}}\|_\infty$ (right plot) as a function of $k \in \{2, 8, 16, 24, 40, 60\}$ with $\gamma \in \{0.2, 0.4, 0.6\}$.

4 Example: A simulation study

The aim of this section is to illustrate how the exact method of Section 2 works in a particular probabilistic error-generating setup.

For a given $k \geq 1$, we consider the simple choice

$$A^* = (\underbrace{I_n \mid I_n \mid \cdots \mid I_n}_{k \text{ times}})^T, \quad b^* = 0_{kn \times 1}.$$

Clearly, the correct solution of $A^*x^* = b^*$ is $x^* = 0$. Now we introduce errors: we run our method with $A^\Delta = E$, $b^\Delta = e$ and

$$A = A^* + U, \quad b = b^* + u,$$

where U is a random $(nk \times n)$-matrix with independent entries sampled from $\mathrm{Unif}(-\gamma, \gamma)$ and u is an $(nk \times 1)$-vector with independent entries sampled again from $\mathrm{Unif}(-\gamma, \gamma)$.

Figure 1 shows:

- average simulated values of δ^{opt} as a function of k for three choices $\gamma \in \{0.2, 0.4, 0.6\}$;
- average simulated values of $\|x^{\mathrm{opt}}\|_\infty$ as a function of k for the same choices of γ.

Figure 1 suggests that the following (not surprising) claims could be true, at least in the particular setup of this simulation study:

- δ^* underestimates γ, but asymptotically δ^* estimates γ consistently;
- asymptotically, the method consistently estimates x^*.

Remark 4. The plotted Figure is for $n = 2$. Additional simulations (not presented here) showed that for other values of n the graphs would confirm similar trends. Of course, it is not surprising that the speed of convergence depends on all involved parameters n, k, γ.

Experiments with other distributions with support $[-\gamma, \gamma]$ confirm analogous behavior, too.

5 Conclusion

We considered the TLS problem with Chebyshev norm and designed an algorithm which solves 2^n generalized linear fractional programs. We proved that the problem is NP-hard and thus it cannot be expected that our method could be significantly improved. But the complexity of our method can be understood as good news, since it is exponential in the number of regression parameters but *not* in the number of observations. (In practice we have usually regression models with a low number of regression parameters compared to the number of observations; only rarely we meet regression model with more that 20 regressors, say.) Then, we designed a heuristic method utilizing efficiently computable lower and upper bounds on the optimal residual value. Finally, we illustrated the behavior of our method by a simulation study.

Acknowledgments

M. Hladík was supported by the Czech Science Foundation Grant P402/13-10660S. M. Černý was supported by the Czech Science Foundation Grant P402/12/G097.

References

[1] Michal Černý, Jaromír Antoch, and Milan Hladík. On the possibilistic approach to linear regression models involving uncertain, indeterminate or interval data. *Inf. Sci.*, 244:26–47, 2013.

[2] M. Fiedler, J. Nedoma, J. Ramík, J. Rohn, and K. Zimmermann. *Linear Optimization Problems with Inexact Data.* Springer, New York, 2006.

[3] R. W. Freund and F. Jarre. An interior-point method for multifractional programs with convex constraints. *J. Optim. Theory Appl.*, 85(1):125–161, 1995.

[4] Laurent El Ghaoui and Hervé Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Anal. Appl.*, 18(4):1035–1064, 1997.

[5] Gene H. Golub and Charles F. Van Loan. An analysis of the total least squares problem. *SIAM J. Numer. Anal.*, 17:883–893, 1980.

[6] Milan Hladík. New operator and method for solving real preconditioned interval linear equations. *SIAM J. Numer. Anal.*, 52(1):194–206, 2014.

[7] Milan Hladík and Michal Černý. Interval data, linear regression and minimum norm estimators: Computational issues, 2014. submitted.

[8] Jaroslav Horáček and Milan Hladík. Computing enclosures of overdetermined interval linear systems. *Reliab. Comput.*, 19(2):142–155, 2013.

[9] Jaroslav Horáček and Milan Hladík. Subsquares approach – a simple scheme for solving overdetermined interval linear systems. In R. Wyrzykowski, J. Dongarra, K. Karczewski, and J. Waśniewski, editors, *Parallel Processing and Applied Mathematics*, volume 8385 of *LNCS*, pages 613–622. Springer, 2014.

[10] D. Jukić, T. Marošević, and R. Scitovski. Discrete total l_p-norm approximation problem for the exponential function. *Appl. Math. Comput.*, 94(2-3):137–143, 1998.

[11] Vladik Kreinovich, Anatoly Lakeyev, Jiří Rohn, and Patrick Kahl. *Computational Complexity and Feasibility of Data Processing and Interval Computations.* Kluwer, 1998.

[12] Ivan Markovsky and Sabine Van Huffel. Overview of total least-squares methods. *Signal Process.*, 87(10):2283–2302, 2007.

[13] Tomislav Marošević. A choice of norm in discrete approximation. *Math. Commun.*, 1(2):147–152, 1996.

[14] Yu. E. Nesterov and A. S. Nemirovskij. An interior-point method for generalized linear-fractional programming. *Math. Program.*, 69(1B):177–204, 1995.

[15] Arnold Neumaier. *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge, 1990.

[16] W. Oettli and W. Prager. Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides. *Numer. Math.*, 6:405–409, 1964.

[17] M. R. Osborne and G. A. Watson. An analysis of the total approximation problem in separable norms, and an algorithm for the total l_1 problem. *SIAM J. Sci. Stat. Comput.*, 6(2):410–424, 1985.

[18] Jiří Rohn. Systems of linear interval equations. *Linear Algebra Appl.*, 126(C):39–78, 1989.

[19] Jiří Rohn. Enclosing solutions of overdetermined systems of linear interval equations. *Reliab. Comput.*, 2(2):167–171, 1996.

[20] Jiří Rohn. A manual of results on interval linear problems. Technical Report 1164, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2012.

[21] Siegfried M. Rump. Verification methods for dense and sparse systems of equations. In Jürgen Herzberger, editor, *Topics in Validated Computations*, Studies in Computational Mathematics, pages 63–136, Amsterdam, 1994. Elsevier.

[22] S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem*. SIAM, 1991.

[23] G. A. Watson. Choice of norms for data fitting and function approximation. *Acta Numer.*, 7:337–377, 1998.

[24] G. A. Watson. Data fitting problems with bounded uncertainties in the data. *SIAM J. Matrix Anal. Appl.*, 22(4):1274–1293, 2001.